Molecular Evolution of the Interphotoreceptor Retinoid Binding Protein (IRBP) in Ground Squirrels and Blind Mole Rats

Research Thesis

Presented in partial fulfillment of the requirements
for graduation with Research Distinction in Biology
in the Undergraduate Colleges of The Ohio State University

by Bethany N. Army

The Ohio State University
December 2015

Project Advisor: Dr. Ryan W. Norris, Department of Evolution, Ecology and Organismal Biology

Abstract

Blind mole rats are subterranean rodents and spend most of their lifetime underground. Because these mammals live in a dark environment, their eyes have physically changed and even the role played by their eyes has changed. The processing of visual information begins in the retina, with the detection of light by photoreceptor cells. Interphotoreceptor retinoid binding protein (IRBP) is a transport protein, which facilitates exchange between photoreceptors and retinal-pigmented epithelium. The IRBP gene has been commonly used in rodent phylogenetic analyses, so I analyzed IRBP in blind mole rats to interpret its evolution and to see if there were any changes among the amino acids in the protein. To do this I used molecular data (cytochrome b and IRBP) from two groups of rodents, the muroids and the marmotines. I constructed phylogenetic trees and interpreted the amino acid sequence of IRBP. Blind mole rats had a slower rate of evolution than expected, but there were changes along the amino acid sequence of IRBP, which may indicate that the IRBP gene is functioning in a specialized way to meet the needs of the particular lifestyle of this mammal.

Introduction:

Blind mole rats, subfamily Spalacinae, are burrowing mammals that spend most of their life underground. Life underground provides protection from predators, extreme environmental conditions, and provides far less temperature fluctuations than are present aboveground (Nevo 2007). In a dark underground environment, vision tends to be unnecessary, so many features of the visual system have been reduced in blind mole rats. Reductions have been described in eye size, ocular muscles, minimized optic nerve, regression in size of brain visual centers and retinal changes in size and structure (Nevo

2007). These individuals are truly blind, and their eyes are completely covered by a layer of skin. Although these mammals are unable to form images, their retina is still able to detect light (Nevo 2013).

The processing of visual information begins in the retina with the detection of light by photoreceptor cells, and the biological conversion of a photon into an electrical signal in the retina is called the visual cycle (Crouch 2009). The visual cycle allows the eye to adapt to environmental challenges (Gonzalez-Fernandez 2002). The purpose of the retina is to receive light, convert the light into neural signals and send neural signals on to the brain for visual recognition. Interphotoreceptor retinoid binding protein (IRBP) is a component of the interphotoreceptor matrix of the retina, and it is important to the function and development of the vertebrate retina (Loew 2002). IRBP is thought to be a transport protein, which aids in protection of 11-cis retinal and all trans retinol while facilitating their exchange between photoreceptors and retinal-pigmented epithelium (Lowe 2002).

The subfamily Spalacinae is restricted to the blind mole rats of southeastern Europe, southwestern Asia, the Mediterranean Middle East and northern Africa (Musser and Carleton 2005). Within the Spalacinae subfamily there are two genera, *Spalax* and *Nannospalax*. Blind mole rats belong to the family Spalacidae in the superfamily Muroidea (Musser and Carleton 2005). The Muroidea is a large superfamily of rodents, which includes hamsters, mice, gerbils, true mice, rats and other relatives. The family Spalacidae contains three subfamilies (Norris in prep), the Spalacinae (blind mole rats), Myospalacinae (zokors) and Rhizomyinae (bamboo rats and root rats). All spalacids are subterranean; they spend most of their life underground. Although Norris et al. (2004)

suggested that all spalacids evolved from a subterranean ancestor, the current thinking is that the subterranean condition in the three subfamilies evolved independently (Flynn 2009).

Nannospalax ehrenbergi was the individual species that I used in my study as a representative of the blind mole rat subfamily, Spalacinae. In my analysis, I included *Myospalax* from the subfamily Myospalacinae, and I included *Rhizomys, Cannomys* and *Tachyoryctes* from the subfamily Rhizomyinae. Members from the genera *Rhizomys, Cannomys* and *Tachyoryctes* tend to forage mostly above ground, and rely on their vision to detect predators (Norris in prep). In contrast, individuals from the genus *Myospalax* have smaller eyes and will only occasionally forage above ground, usually at night. None have eyes that are as reduced as the Spalacinae. I also included representatives of all other muroid families (Appendix 1).

A fossorial lifestyle can also be found in many other rodent groups. An example of these is the tribe Marmotini in the family Sciuridae. The Marmotini includes both the chipmunks and the Holarctic ground squirrels (Patterson and Norris in press). Holarctic ground squirrrels build burrows, but spend much of their lives above ground.

The IRBP gene has been commonly used in rodent phylogenetic analyses including work on muroids (Schenk et al. 2013) and marmotines (Patterson and Norris in press). The blind mole rats have reductions in their visual system, so I analyzed the rate of molecular evolution in *Nannospalax ehrenbergi* to see if there were any changes in amino acids of the IRBP protein compared to its fossorial relatives, other members of the Muroidea, and another group of fossorial rodents (ground squirrels). Because they are

unable to form images on their retinas, I hypothesized that the IRBP gene would evolve more rapidly in blind mole rats due to relaxed selective pressure.

Methods

Calibrating Molecular Clock

Molecular clocks require fossils to calibrate them. Fossils can only be assigned with certainty when used as minimum values, because there is always a possibility that an older fossil could be discovered, and that would change the results. Norris et al. (2015) developed two related methods (PenG and GLin) that use information from the oldest known fossil and immediately younger fossils to create a prior probability distribution for the true age of a node. They demonstrated the effectiveness of the methods using simulated fossil and genetic data. Unfortunately, the PenG and GLin methods were developed using simulated fossils that are known with complete precision. In real datasets, the dating of fossils frequently comes with uncertainty. For example, an individual fossil might be dated to 10 to 14 million years ago (Mya) instead of exactly 12 Mya. So, to deal with imprecisely dated fossils, I started by creating a lognormal prior in BEAUti (Drummond et al. 2013) based on the PenG or GLin method using the median values of relevant intervals and recorded the median and upper 95% quantile of the distribution. This prior was discarded, and the actual prior on the node was set using the minimum possible value of the oldest fossil as the zero offset. The median was set to match the median of the discarded prior. Finally, the upper 95% quantile was forced to match the discarded upper 95% quantile by adjusting the standard deviation of the distribution.

The details of fossils used as calibrations are shown in Appendix 2. Five calibrations were applied to the muroid dataset. A single fossil calibration was applied to the marmotine dataset.

An additional calibration has been suggested for marmotines based on the possible presence of Asian *Eutamias* chipmunk fossils from the Chattian (23-26 Mya; Meng et al. 2008; Patterson and Norris in press). These fossils may not be completely accurate because they are identified as *Eutamias* indet. instead of confidently assigned to the genus. Meanwhile, the genus *Eutamias* has only recently been used to refer to only one species, *Eutamias sibiricus*. The genus *Eutamias* was once applied to any chipmunk where a third upper premolar (P3) was present (Patterson and Norris in press). Now, molecular data indicate that chipmunks with P3 do not form a monophyletic group; therefore the possibility exists that this Chattian fossil might be the result of an earlier migration into Asia from North America. Because this Asian chipmunk fossil cannot clearly be tied to the modern Siberian chipmunk, *Eutamias sibiricus,* I excluded this potential calibration. Instead, I used this study to test whether the timing of this fossil is consistent with what is known about marmotine divergence times.

Analyses

All sequences were downloaded from GenBank. The Marmotini cytochrome b (cytb) and IRBP sequences and alignments came from the dataset of Patterson and Norris (in press). The Muroid IRBP sequences and alignments were obtained from the dataset of Schenk et al. (2013). There were a total of 49 species included in my analyses, 27 of which were muroids and 22 of which were marmotines (Appendix 1).

For the muroid analysis, the appropriate model of evolution was determined by using jmodeltest (version 2.1.7; Darriba et al. 2012). The models of evolution used for cytb and IRBP were both GTR+I+G. Phylogenetic trees were constructed under a Bayesian analysis using BEAST (version 1.8.2; Drummond et al. 2012). Across the two partitions, trees were linked but substitution models and clock models were not linked. The tree was constrained to force monophyly at all nodes that were well supported (Bayesian posterior probability [PP] > 0.95 and maximum likelihood bootstrap > 90%) in Schenk et al. (2013). I used a lognormal relaxed molecular clock and a Yule speciation model. Fossil calibrations were used as priors as indicated in Appendix 2. The Markov Chain Monte Carlo (MCMC) was run for 30 million generations sampling every 3000 generations. A burnin of 1000 trees was applied after visualizing the runs in Tracer (version 1.6.0; Drummond et al. 2012).

For the Marmotini analysis, the models of evolution used for cytb was HKY+G and for IRBP was GTR+G. Across the two partitions, trees were linked but substitution models and clock models were not linked. No constraints on tree topology were applied. A single calibration was used as described in Appendix 2. Other parameters were the same as for the muroid dataset.

To visualize changes in the IRBP protein, I translated the nucleotides into amino acids using Mesquite (3.02; Maddison and Maddison 2015). Unique changes in the amino acid sequences for *Nannospalax ehrenbergi* were visualized using Mesquite.

Results

Figure 1 shows the results of the muroid analysis. The Spalacidae diverged from the other Muroidea 38.0 Mya (95% highest posterior density [HPD] = 32.9 Mya to 43.4 Mya). The blind mole rats split from *Myospalax* around 28 Mya (95% HPD = 24.5 Mya to 32.8 Mya), and the rhizomyines split from the blind mole rats around 26 Mya (95% HPD = 24.0 Mya to 30.2 Mya). The estimated rate of evolution of IRBP along the *Nannospalax ehrenbergi* branch is 0.0023 (95% HPD = 0.0016 to 0.0031), whereas the mean rate of evolution of IRBP across all muroids is 0.0026 (95% HPD = 0.0023 to 0.0033). The estimated rate of evolution of cytb along the *Nannospalax ehrenbergi* branch is 0.028 (95% HPD = 0.016 to 0.038), whereas the mean rate of evolution of cytb across all muroids is 0.036 (95% HPD= 0.028 to 0.044).

Figure 2 shows the results of the marmotine analysis. The chipmunks diverged from other marmotines 28.4 Mya (95% HPD = 26.3 Mya to 38.0). *Eutamias* and *Tamias* split from one another 11.8 Mya (95% HPD = 7.3 Mya to 17.1 Mya); these values indicate the split was after the Chattian fossil (23.8 Mya to 26 Mya). The mean rate of IRBP evolution within the marmotines is 0.0011 (95% HPD = 0.00075 to 0.0014). The mean rate of cytb evolution within the marmotines is 0.050 (95% HPD = 0.032 to 0.070). The ratio of the mean IRBP rate to the mean cytb rate in *Nannospalax ehrenbergi is* 0.082, the ratio in muroids is 0.072, and the ratio in marmotines is 0.022.

Table 1 shows amino acid changes unique to *Nannospalax ehrenbergi*. There are seven unique substitutions, and a deletion of four amino acids found only in the blind mole rats. All other muroids and all marmotines have identical amino acids at these sites except for in position 91, where Holarctic ground squirrels (subtribe Marmotina) have

glutamine (Q) instead of arginine (R). The blind mole rat sample exhibits a deletion of a four amino acid sequence, valine-serine-arginine-serine (VSRS), at positions 239 to 242 (numbered based on Appendix 3). All other muroids and all marmotines showed an identical amino acid sequence (VSRS) in this region.

Discussion

The rate of change in the IRBP protein in blind mole rats is surprisingly slow compared to other muroids (Fig. 1). This rate was 0.0023 compared to 0.0026 (95% HPD = 0.0023 to 0.0033) among the muroids as a whole. Although it is not outside the 95% HPD for muroids, it is still among the slowest evolving branches in muroids. These results were very different from what I first hypothesized, which is that the IRBP gene would evolve faster in blind mole rats due to their much-reduced visual system. Amino acid changes did occur, including several unique changes (Table 1), but there were relatively few and no stop codons were present. This suggests that the IRBP gene must still have an important function in blind mole rats. Even though the blind mole rats have reductions in eye size, ocular muscles, a minimized optic nerve, regression in the size of brain visual centers, retinal changes in size and structure, and other changes (Nevo 2007); they appear to still use IRBP for some sort of function. Even though the IRBP rate of evolution in blind mole rats is slow, the ratio of the mean IRBP rate to the mean cytb rate in *Nannospalax ehrenbergi is* 0.082, which is generally high compared to the other ratios.

The amino acid changes that were present in the blind mole rat were completely unique to that species (Table 1). Because of these seven unique changes and a deletion of four amino acids (VSRS), this leads me to believe that they might be using IRBP strictly

for a specialized light detection for their own needs, compared to the other species with a normally functioning retina. The mean rate of IRBP evolution within the marmotines is 0.0011, which is lower than the blind mole rats and much lower than muroids overall. In order to test the true function of IRBP in *Nannospalax ehrenbergi,* other tests would have to be applied. For example, researchers could benefit by studying knockout laboratory mice or mice whose IRBP has been replaced with the IRBP from *Nannospalax ehrenbergi*. Other possibilities might be to run a more detailed investigation of IRBP active sites, sequence IRBP from more members of the Spalacinae, sequence other genes that IRBP interacts with, or perform *in silico* analyses to determine the protein's folding actions.

The analyses that I ran incorporated both PenG and GLin methods (Norris et al. 2015), which use information from the oldest known fossil and immediately younger fossils to create a prior probability distribution for the true age of a node. These methods were just published this year, and this is the first time that they have been applied to real data sets. Modifications had to be applied due to the fact that many of the fossils could not be expressed as a single age, but were dated to an age interval. To solve this problem, we generated estimates based on the median value of the age intervals of the two oldest known fossils. Our solution yielded a sensible value for zero offset, median, and upper 95%, but the standard deviation curve was essentially arbitrary.

My results (Fig. 1) suggest a divergence time between Spalacidae and other muroids at 38.0 Mya (95% HPD = 32.9 Mya to 43.4 Mya), compared to Schenk et al. (2013) who estimated this age as roughly 37 Mya (95% HPD ~ 32 Mya to 40 Mya based on their Fig. 4). In my analysis the divergence times are comparable to Schenk et al. (2013), but my

95% HPD range is greater than Schenk et al. I expected these results because I used the PenG and GLin methods. These methods are expected to have a large 95% HPD, because they incorporate the possibility that older fossils do exist but they have not yet been discovered (Norris et al. 2015).

The Spalacidae is usually treated as a family that combines blind mole rats, zokors, bamboo rats, and root rats (Norris et al. 2004; Musser and Carleton 2005; Flynn 2009; Schenk et al. 2013; Norris in prep.) My results suggest that the three subfamilies of Spalacidae diverged in the Late Oligocene and these subfamilies appear to be as old or older than other families within the Muroidea. For example, Cricetidae and Muridae diverged about 25 Mya (95% HPD = 20.7 Mya to 29.3 Mya), which may be more recent than the split between Spalacinae and Rhizomyinae (26 Mya; 95% HPD = 24.0 Mya to 30.2 Mya).

A similar argument could be made with the family Dipodidae. Holden and Musser (2005) treated birch mice (*Sicista*), jumping mice (*Zapus* and *Napaeozapus*), and jerboas as part of a single family. My results suggest that birch mice diverged 35.4 Mya (95% HPD = 27.9 Mya to 42.7 Mya) and that jumping mice diverged from jerboas 28.4 Mya (95% HPD = 22.0 Mya to 35.3 Mya). The results shown in Figure 1 imply that the Dipodidae family could be separated into three different families. Therefore, *Jaculus, Dipus* and *Allactaga* could be members of the Dipodidae, *Zapus* and *Napaeozapus* could be included in a Zapodidae family and *Sicista* could be included in a Sicistidae family. This arrangement of three families in one superfamily (Dipodoidea) has been used in the past (Holden and Musser 2005).

For the Marmotini dataset, my results agree with the analysis of Patterson and Norris (in press). This is not surprising since the datasets largely overlap. Patterson and Norris used uncalibrated phylogenetic trees; whereas my data included fossil-based uncertainty. The combined cytb and IRBP dataset provides evidence that the three chipmunk lineages are distinct and should be recognized as separate genera. *Neotamias* diverged from the other chipmunks 17.1 Mya (95% HPD = 11.1 Mya to 24.6 Mya) and *Eutamias* diverged from *Tamias* 11.8 Mya (95% HPD = 7.3 Mya to 17.1 Mya). This is older than or comparable to divergence times among genera of Holarctic ground squirrels. For example, *Ictidomys* and *Poliocitellus* diverged 5.8 Mya (95% HPD = 3.6 Mya to 8.9 Mya). In the Muroidea, even members of different tribes appear to have diverged after the divergence among chipmunks. Examples are *Tachyoryctes* and *Rhizomys* at 12.3 Mya (95% HPD = 9.5 Mya to 15.7 Mya) and *Arvicanthis* and *Mus* at 13.4 Mya (95%HPD = 11.2 to 17.0 Mya).

The GLin and PenG approaches appear to be effective new methods that can be adapted to challenges present in real fossil datasets. Cytochrome b and IRBP were effective genes in evaluating molecular evolution in muroid and marmotine rodents. Although the evolution rate of IRBP in blind mole rats was unexpectedly slow, there are amino acid changes and deletions within the protein that suggests that IRBP may be functioning in a different way than in other muroids and marmotines.

Acknowledgements

I would like to thank Bruce D. Patterson for assisting the Norris lab with the ground squirrel and chipmunk datasets. I would also like to thank Jorgi Schramm and

Hannah Downer for assisting me in finding fossil calibrations. Lastly, I would like to thank Dr. Kameswarrao Casukhela for assistance with the determining the shape of the Bayesian priors for the fossil calibrations.

References

Crouch, R. K. 2009. The Visual Cycle. Photobiological Sciences Online (PSO). http://www.photobiology.info/Crouch.html

Darriba, D., G. L. Taboada R. Doallo, and D. Posada. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9:772.

Drummond A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012 Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology And Evolution 29:1969-1973.

Flynn, L. J. 2009. The antiquity of *Rhizomys* and independent acquisition of fossorial traits in subterranean muroids. Bulletin of the American Museum of Natural History 331:128-156.

Gonzalez-Fernandez, F. 2002. Evolution of the Visual Cycle: The Role of Retinoid-binding Proteins. Journal of Endocrinology 175:75-88.

Holden, M. E. and G. G. Musser. 2005. Family Dipodidae. Pp 871-893 in Mammal Species of the World (D. E. Wilson and D. M. Reeder eds.). The Johns Hopkins University Press.

Kimura, Y. M. T. R. Hawkins, M. M. Mcdonough, L. L. Jacobs, and L. J. Flynn. 2015. Corrected placement of *Mus - Rattus* fossil calibration forces precision in the molecular tree of rodents. Scientific Reports 5:14444.

Loew, A. and F. Gonzalez-Fernandez. 2002. Crystal Structure of the Functional Unit of Interphotoreceptor Retinoid Binding Protein. Structure 10:43-49.

López-Antoñanzas, R. 2009. First *Potwarmus* from the Miocene of Saudi Arabia and the early phylogeny of murines (Rodentia: Muroidea). Zoological Journal of the Linnean Society 156:664-679.

López-Antoñanzas, R., F. Knoll, S. Wan, and L. J. Flynn. 2015. Causal evidence between monsoon and evolution of rhizomyine rodents. Scientific Reports 5:9008.

Maddison, W. P. and D. R. Maddison. 2015. Mesquite: a modular system for evolutionary analysis. Version 3.04 http://mesquiteproject.org

Meng, J., J. Ye, W.-Y. Wu, X.-J. Ni, and S.-D. Bi. 2008. The Neogene Dingshanyanchi Formation in Northern Jungguar basin of Xinjiang and its stratigraphic implications. Vertebrata PalAsiatica 46: 90–110.

Musser, G. G. and M. D. Carleton. 2005. Superfamily Muroidea. Pp 894-1531 in Mammal Species of the World (D. E. Wilson and D. M. Reeder eds.). The Johns Hopkins University Press.

Nevo, E. 2007. Mosaic Evolution of Subterranean Mammals: Tinkering, Regression, Progression, and Global Convergence. Pp 375-388 in Subterranean Rodents: News from Underground (S. Begall, H. Burda, and C. E. Schleich eds.). Springer.

Nevo, E. 2013. Stress, Adaptation, and Speciation in the Evolution of the Blind Mole Rat, Spalax, in Israel. Molecular Phylogenetics and Evolution 66.2: 515-25.

Norris, R. W. In preparation. Family Spalacidae. In Handbook of the Mammals of the World - Volume 7 Rodents II (D. E. Wilson and R. A. Mittermeier eds.). Lynx Edicions.

Norris, R. W., C. L. Strope, D. M. McCandlish, and A. Stoltzfus. 2015. Bayesian Priors for Tree Calibration: Evaluating Two New Approaches Based on Fossil Intervals. bioRxiv doi: http://dx.doi.org/10.1101/014340.

Norris, R. W., K. Zhou, C. Zhou, G. Yang, C. William Kilpatrick, and R. L. Honeycutt. 2004. The Phylogenetic Position of the Zokors (Myospalacinae) and Comments on the Families of Muroids (Rodentia). Molecular Phylogenetics and Evolution 31:972-78.

Paleobiology Database (Fossilworks.org). Last accessed 24 November 2015.

Patterson B.D. and R.W. Norris. In press. Towards a uniform nomenclature for ground squirrels:the status of the Holarctic chipmunks. Mammalia. DOI 10.1515/mammalia-2015-0004.

Schenk, J. J., K. C. Rowe, and S. J. Steppan. 2013. Ecological Opportunity and Incumbency in the Diversification of Repeated Continental Colonizations by Muroid Rodents. Systematic Biology 62: 837-64.

Table 1. Amino Acid changes unique to *Nannospalax ehrenbergi*. Position number corresponds to Appendix 3.

Position #	**Amino Acid Change**
3	G→A
42	R→C
91	R→W
127	R→W
150	V→I
154	D→B
246	L→M
239-242	VSRS → Deletion

Figure 1. Time calibrated maximum clade credibility tree from BEAST for the muroid dataset. Error bars represent 95% highest posterior density (HPD). Line thickness represents rate of evolution of IRBP gene. Thicker lines indicate faster rate.

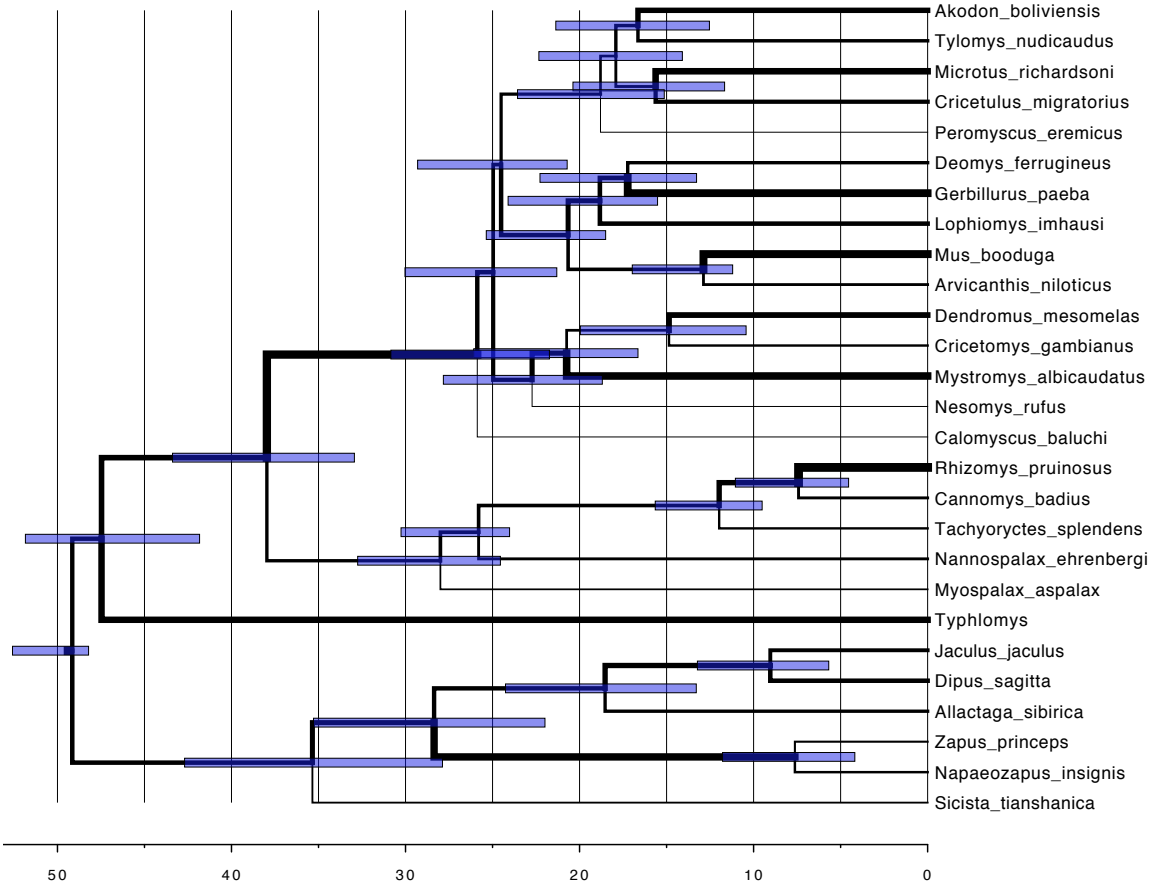
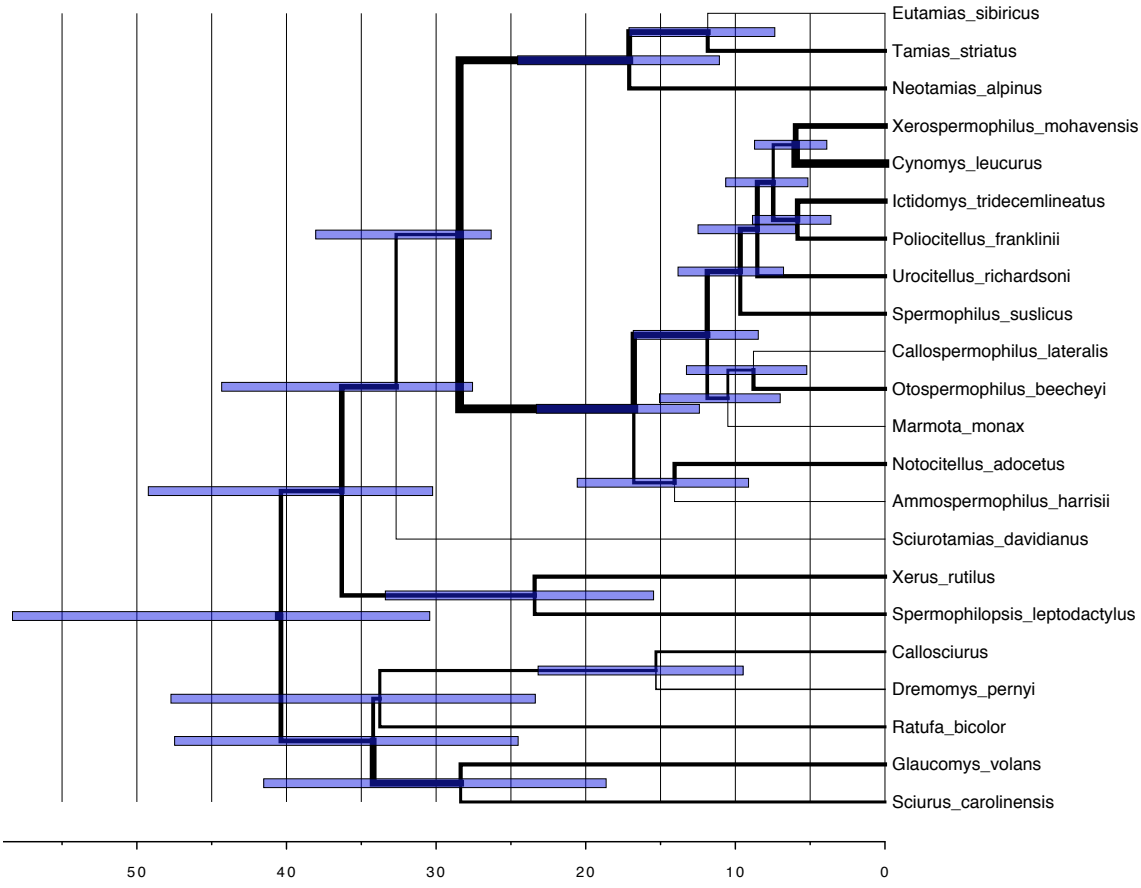


Figure 2. Time calibrated maximum clade credibility tree from BEAST for the marmotine dataset. Error bars represent 95% highest posterior density (HPD). Line thickness represents rate of evolution of IRBP gene. Thicker lines indicate faster rate.



Appendix 1. GenBank numbers and lifestyle for taxa used in phylogenetic analyses. Taxonomy based on Musser and Carleton (2005) and Norris (in prep.) for Muroidea, Holder and Musser (2005) for Dipodoidea, and Patterson and Norris (in press) for Sciuridae. 'Subterranean' represents animals adapted to life underground and 'fossorial' represents animals that construct extensive burrow systems, but spend considerable time aboveground. The category 'terrestrial' includes arboreal animals in this table.

Taxon	Lifestyle	IRBP	CYTB
Muroidea			
Spalacidae			
Spalacinae			
Nannospalax ehrenbergi	subterranean	KC953405	JX451832
Myospalacinae			
Myospalax aspalax	subterranean	AY326097	AF326272
Rhizomyinae			
Rhyzomyini			
Cannomys badius	subterranean	KC953363	—
Rhizomys pruinosus	fossorial	AF297283	NC021478
Tachyoryctini			
Tachyoryctes splendens	subterranean	AY326112	AF160602
Calomyscidae			
Calomyscus baluchi	terrestrial	AY163581	EU135591
Cricetidae			
Arvicolinae			
Microtus richardsoni	semi-fossorial	AM919404	AF163905
Cricetinae			
Cricetulus migratorius	semi-fossorial	KC953367	AY288508
Neotominae			
Peromyscus eremicus	terrestrial	KC953436	DQ973100
Sigmodontinae			
Akodon boliviensis	semi-fossorial	KC953351	KC841367
Tylomyinae			
Tylomys nudicaudus	terrestrial	AY163643	DQ179812
Muridae			
Deomyinae			
Deomys ferrugineus	terrestrial	KC953373	FJ415478
Gerbillinae			
Gerbillurus paeba	semi-fossorial	KC953376	JF704124
Lophiomyinae			
Lophiomys imhausi	terrestrial	KC953389	KR089025
Murinae			
Arvicanthini			
Arvicanthis niloticus	semi-fossorial	DQ022386	AF004572
Murini			
Mus booduga	semi-fossorial	AB125796	AB125761

Taxon	Lifestyle	IRBP	CYTB
Muroidea (continued)			
Nesomyidae			
Dendromurinae			
Dendromus mesomelas	terrestrial	KC853371	KF811233
Cricetomyinae			
Cricetomys gambianus	semi-fossorial	KC953366	AF160614
Mystromyinae			
Mystromys albicaudatus	semi-fossorial	AY163594	AF160607
Nesomyinae			
Nesomys rufus	semi-fossorial	AY326099	AF160592
Platacanthomyidae			
Typhlomys	terrestrial	GQ272606	KC209557
Dipodoidea			
Dipodidae			
Allactaginae			
Allactaga sibirica	semi-fossorial	AY326076	XJ0309SD03
Dipodinae			
Dipus sagitta	semi-fossorial	AJ427232	AM407909
Jaculus jaculus	semi-fossorial	AM407907	JN214545
Sicistinae			
Sicista tianshanica	semi-fossorial	AF297288	KM397204
Zapodinae			
Napaeozapus insignis	semi-fossorial	AY326098	AJ389535
Zapus princeps	semi-fossorial	AF297287	KF441295
Sciuridae			
Xerinae			
Marmotini			
Tamiina			
Tamias striatus	semi-fossorial	JN414824	JN042555
Eutamias sibiricus	semi-fossorial	AB253981	KF990333
Neotamias alpinus	semi-fossorial	—	KJ452914
Marmotina			
Ammospermophilus harrisii	fossorial	AF157926	AF157926
Callospermophilus lateralis	semi-fossorial	AY227586	AF157950
Cynomys leucurus	fossorial	AY227584	JQ885590
Ictidomys tridecemlineatus	fossorial	AF297278	KP698974
Marmota monax	fossorial	AJ427237	AF157953
Notocitellus adocetus	fossorial	—	AF157843
Otospermophilus beecheyi	semi-fossorial	—	AF157918
Poliocitellus franklinii	fossorial	—	AF157894
Spermophilus suslicus	fossorial	—	AF157897
Urocitellus richardsoni	fossorial	—	AF157914
Xerospermophilus mohavensis	fossorial	JX065593	AF157928

Taxon	Lifestyle	IRBP	CYTB
Sciuridae (continued)			
Xerinae (continued)			
Xerini			
Spermophilopsis leptodactylus	fossorial	AY227624	AF157865
Xerus rutilus	fossorial	AY227625	AY452690
Tribe *incertae sedis*			
Sciurotamias davidianus	semi-fossorial	AY227621	KC005710
Callosciurinae			
Callosciurus	terrestrial	HQ698524	AB499914
Dremomys pernyi	terrestrial	HQ698525	HQ698363
Ratufa bicolor	terrestrial	AY227608	NC023780
Sciurinae			
Glaucomys volans	terrestrial	AY227598	AF157921
Sciurus carolinensis	terrestrial	HG962385	FJ200744

Appendix 2. Explanation of fossils used to calibrate molecular clock analyses in Fig. 1 and Fig. 2.

Figure 1, node Myodonta: the *Nannospalax – Dipus* divergence. *Elymys* represents the oldest fossil in this clade (Paleobiology Database [PBDB] no. 16218; Schenk et al. 2013). It is a member of the superfamily Dipodoidea and dates to the early Bridgerian (~48.2-50.3 Mya; median = 49.3 Mya). The oldest member of the sister group, the superfamily Muroidea, is *Pappocricetodon* from ~45 Mya (PBDB 37493). I used the GLin method (Norris et al. 2015) to initially determine the settings for a lognormal prior based on what median values for these fossils would look like using BEAUti (zero offset = 49.3 Mya; median = 51.45 Mya; stdev = 1.814; upper 95% = 91.77 Mya). The actual prior used in the analysis was set so that the zero offset matched the minimum value of the *Elymys* fossil (48.2 Mya), the median and upper 95% matched the values from the initial estimation (median = 51.45 Mya; upper 95% = 91.77 Mya), and the standard deviation was adjusted to generate this median and this upper 95%. The standard deviation was set to 1.5596.

Figure 1, node Spalacinae + Rhizomyinae: the *Nannospalax – Rhizomys* divergence. *Prokanisamys* represents the oldest fossil in this clade (Flynn 2009). It is a member of the subfamily Rhizomyinae and dates to 24-27 Mya median = 25.5 Mya). The oldest member of the sister group, the lineage including Spalacinae, is *Eumyarion* from ~23-24 Mya (Flynn 2009). I used the GLin method (Norris et al. 2015) to initially determine the settings for a lognormal prior based on what median values for these fossils would look like using BEAUti (zero offset = 25.5 Mya; median = 26.5 Mya; stdev = 1.814; upper 95% = 45.26 Mya). The actual prior used in the analysis was set so that the zero offset matched the minimum value of the *Prokanisamys* fossil (24 Mya), the median and upper 95% matched the values from the initial estimation (median = 26.5 Mya; upper 95% = 45.26 Mya), and the standard deviation was adjusted to generate this median and this upper 95%. The standard deviation was set to 1.3015.

Figure 1, node Rhizomyinae: the *Rhizomys – Tachyoryctes* divergence. *Miorhizomys* represents the oldest fossil in this clade (López-Antoñanzas et al. 2015). It is a member of the tribe Rhizomyini and dates to ~9.5 Mya. The oldest member of the sister group, the tribe Tachyoryctini, is *Tachyoryctes pliocaenicus* from ~3.5 Mya (López-Antoñanzas et al. 2015). I used the GLin method (Norris et al. 2015) to set a lognormal prior for the final analysis (zero offset = 9.5 Mya; median = 12.75 Mya; stdev = 1.814; upper 95% = 73.75 Mya).

Figure 1, node Muridae: the *Mus – Gerbillurus* divergence. *Potwarmus* represents the oldest fossil in this clade, and dates to ~18.5 Mya (López-Antoñanzas 2009). The second oldest member is *'Myocricetodon' sivalensis* from ~17 Mya (López-Antoñanzas 2009). I used the PenG method (Norris et al. 2015) to set a lognormal prior for the final analysis (zero offset = 18.5 Mya; median = 20 Mya; stdev = 1.814; upper 95% = 48.13 Mya).

Figure 1, node Murini + Arvicanthini: the *Mus – Arvicanthis* divergence. *"?Karnimata"* represents the oldest fossil in this clade and dates to ~11.2 Ma (Kimura et al. 2015). The second oldest member from an independent locality is *Progonomys* from ~9.2 Mya (PBDB 41922). I used the PenG method (Norris et al. 2015) to set a lognormal

prior for the final analysis (zero offset = 11.2 Mya; median = 13.2 Mya; stdev = 1.814; upper 95% = 50.72 Mya).

Figure 2, node Marmotini: the *Tamias – Marmota* divergence. *Nototamias* represents the oldest fossil in this clade (PBDB 17709; Patterson and Norris in press) and dates to the Geringian (26.3-30.8 Mya; median = 28.55 Mya). The second oldest fossil is also *Nototamias* (PBDB 41506) and dates to the Monroecreekian (24.8-26.3 Mya; median = 25.55 Mya). I used the PenG method (Norris et al. 2015) to initially determine the settings for a lognormal prior based on what median values for these fossils would look like using BEAUti (zero offset = 28.55 Mya; median = 31.55 Mya; stdev = 1.814; upper 95% = 87.86 Mya). The actual prior used in the analysis was set so that the zero offset matched the minimum value of the older *Nototamias* fossil (26.3 Mya), the median and upper 95% matched the values from the initial estimation (median = 31.55 Mya; upper 95% = 87.86 Mya), and the standard deviation was adjusted to generate this median and this upper 95%. The standard deviation was set to 1.4742.

Appendix 3. Amino acid sequence of IRBP gene (KC953405) for *Nannospalax ehrenbergi* after alignment with other samples in the muroid dataset. Numbers used in Table 1 and text correspond to position in this sequence.

LMAMQAAIEQAMKSREILGISDPQTLAH[M/V]LTAGVQSSLNDPCLFISYEPS
TLEAPQ-QAPTLTNFTQEELLAQLQRNIHHEVLEGN[L/V]GYLWVDDLLGQEV
LSKLGEFLVAHMWRQLMNTSALVLDLWHCTGGHVSGIPYVISYLHPGNTI
MHVBTIYDRPSNTTTEIWTLPKVLGERYSADKDVVVLTSGHTGGVAEDIA
YILKQMRRAIMVGEQTEGGALDLQKLRIGQSNFFLTVP----LGPMGGGG
QTWE